Bayesian Reconstruction of Particle Beam Phase Space from Low Dimensional Data

Charles W. Nakhleh, Dave Higdon, Christopher K. Allen, and Vidya Kumar

Los Alamos National Laboratory

1 Introduction

Understanding the physics of charged particle beams (e.g., proton beams) is essential to designing and controlling efficient particle accelerators. The dynamics of a beam is naturally formulated in a six-dimensional phase space (three position and three momentum or velocity dimensions). However, experimental beam profile data taken from accelerators are typically one-dimensional (coordinate) projections of the phase space distribution or image. The objective of this study is to apply Bayesian image reconstruction techniques to reconstruct the phase space images at different points along a beam from a series of one dimensional projection data sets (wirescans). The resulting phase images along the beamline provide key beam diagnostic information and are of great interest to accelerator designers. Because the cardinality of the phase images is much larger than the cardinality of the projection data, successful reconstruction of the phase density depends critically on the incorporation of non-trivial prior information.

The outline of this paper is as follows. First, we describe the specific example used in this paper: proton beam data taken on the Low Energy Demonstration Accelerator (LEDA) at the Los Alamos National Laboratory (Allen, Chan, Colestock, Crandall, Garnett, Gilpatrick, Lysenko, Qiang, Schneider, Schulze, Sheffield, Smith and Wangler (2002)). We then describe the technique of beam transfer matrices used in this paper to simulate the beam propagation. This section includes a comparison of the transfer matrix method with a detailed accelerator simulation code. Then we present the statistical formulation of the phase space reconstruction problem, including the likelihood model; the image prior; and the Markov Chain Monte Carlo (MCMC) technique used to estimate the posterior. Finally, we describe our results and outline additional sensitivity studies we have undertaken.

2 LEDA experiments

LEDA was an 11m long, 6.7 MeV proton accelerator designed specifically to study continuous, high-current proton beams. The LEDA lattice consisted of 52 focusing/defocusing (FODO) quadrupole magnets; a number of steering magnets; and nine wire scanners (Fig. 1). A picture of the actual machine is given in Fig. 2.

The data utilized in this paper come from the middle four (20, 22, 24, and 26) and the last four wirescanners (45, 47, 49, and 51) in the beamline. At each wire scan station, the proton beam distribution was projected onto both the vertical and horizontal directions. The wire scanner (Fig. 3) used a 33 μm carbon wire to measure the beam profile via secondary electron emission (because
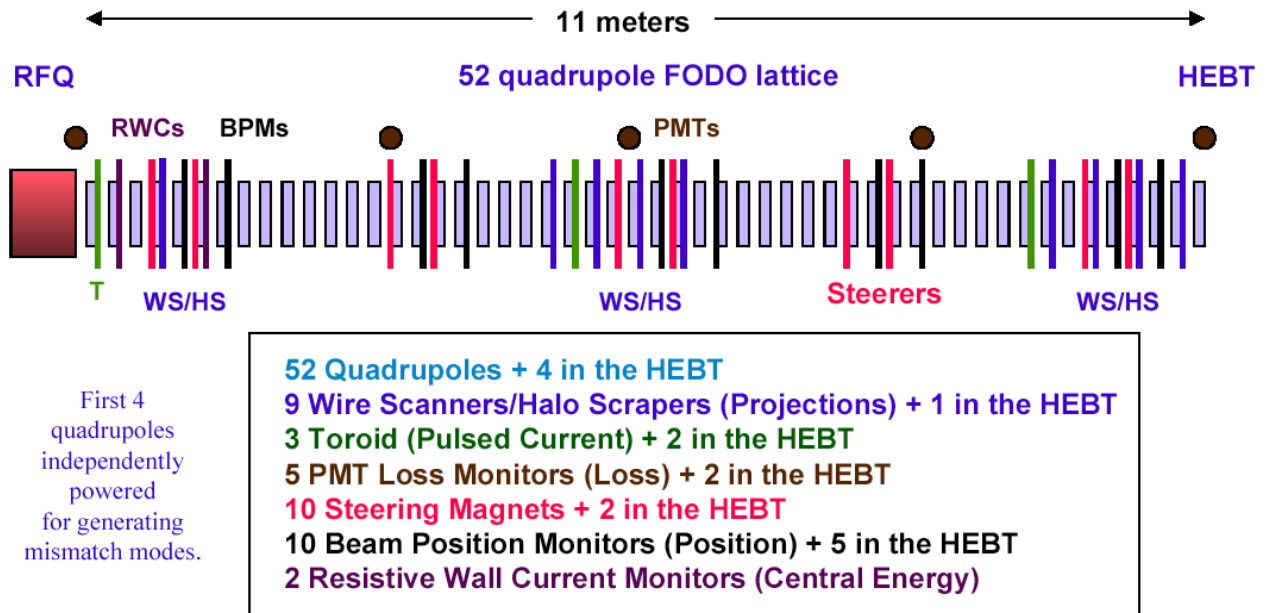
11 meters

RFQ   **52 quadrupole FODO lattice**   **HEBT**

RWCs   **BPMs**   **PMTs**

T   **WS/HS**   **WS/HS**   **Steerers**   **WS/HS**

First 4 quadrupoles independently powered for generating mismatch modes.

52 Quadrupoles + 4 in the HEBT
9 Wire Scanners/Halo Scrapers (Projections) + 1 in the HEBT
3 Toroid (Pulsed Current) + 2 in the HEBT
5 PMT Loss Monitors (Loss) + 2 in the HEBT
10 Steering Magnets + 2 in the HEBT
10 Beam Position Monitors (Position) + 5 in the HEBT
2 Resistive Wall Current Monitors (Central Energy)

Figure 1: The Low Energy Demonstration Accelerator.
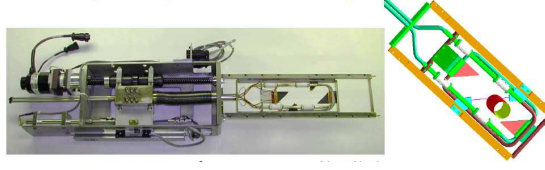

Figure 2: The LEDA experimental lattice.

3

Figure 3: A LEDA wirescanner.

the wire itself was not thick enough to stop the protons themselves). A simulated beam with the associated x direction phase images is shown in Fig. 4.

3 Beam transfer matrices

As the beam travels along its design trajectory, each charged particle in the beam is fully described by its phase space coordinates $\{x(s), x'(s), y(s), y'(s), z(s), z'(x)\}$, where s is the coordinate along the design trajectory, and $x'(s) \equiv dx/ds = p_x/p = \frac{dx/dt}{ds/dt}$, etc., is the relative momentum variable commonly used in beam design (Allen and Pattengale (2002)). We will take x and y to be the transverse directions for the linear LEDA array. The beam particles are subject to both applied external forces from the steering magnets as well as self forces due to space charge effects. Using the notation of Allen and Pattengale (2002), the quadrupole magnets are designed to yield a linear force law:

$$\mathbf{F}_{\text{ext}} = -\mathbf{K}(s) \cdot \mathbf{r}, \tag{1}$$

where $\mathbf{K}(s) = (\kappa_x(s), \kappa_y(s), \kappa_z(s))$ represents the action of the external magnets along the beamline. Note that κ_x can be either positive or negative depending on whether the magnet is focusing or defocusing, respectively, in the x plane, and similarly in the y and z planes.

If we ignore space charge effects, and keep in mind that 6.7 MeV protons are slowly moving and therefore nonrelativistic, the equations of motion of the particle beam are easily written down (Allen and Pattengale, 2002, Eqn. 16):

$$x'' + k_x^2 x = 0, \tag{2}$$

$$y'' + k_y^2 y = 0, \tag{3}$$

$$z'' + k_z^2 z = 0, \tag{4}$$

where $k_x^2 = \kappa_x/m\mathbf{v}^2$, and $k_x(s)$ is a function of position along the beam. Despite appearances, k_x^2 can be either positive (focusing magnet) or negative (defocusing magnet), and similarly for k_y^2.
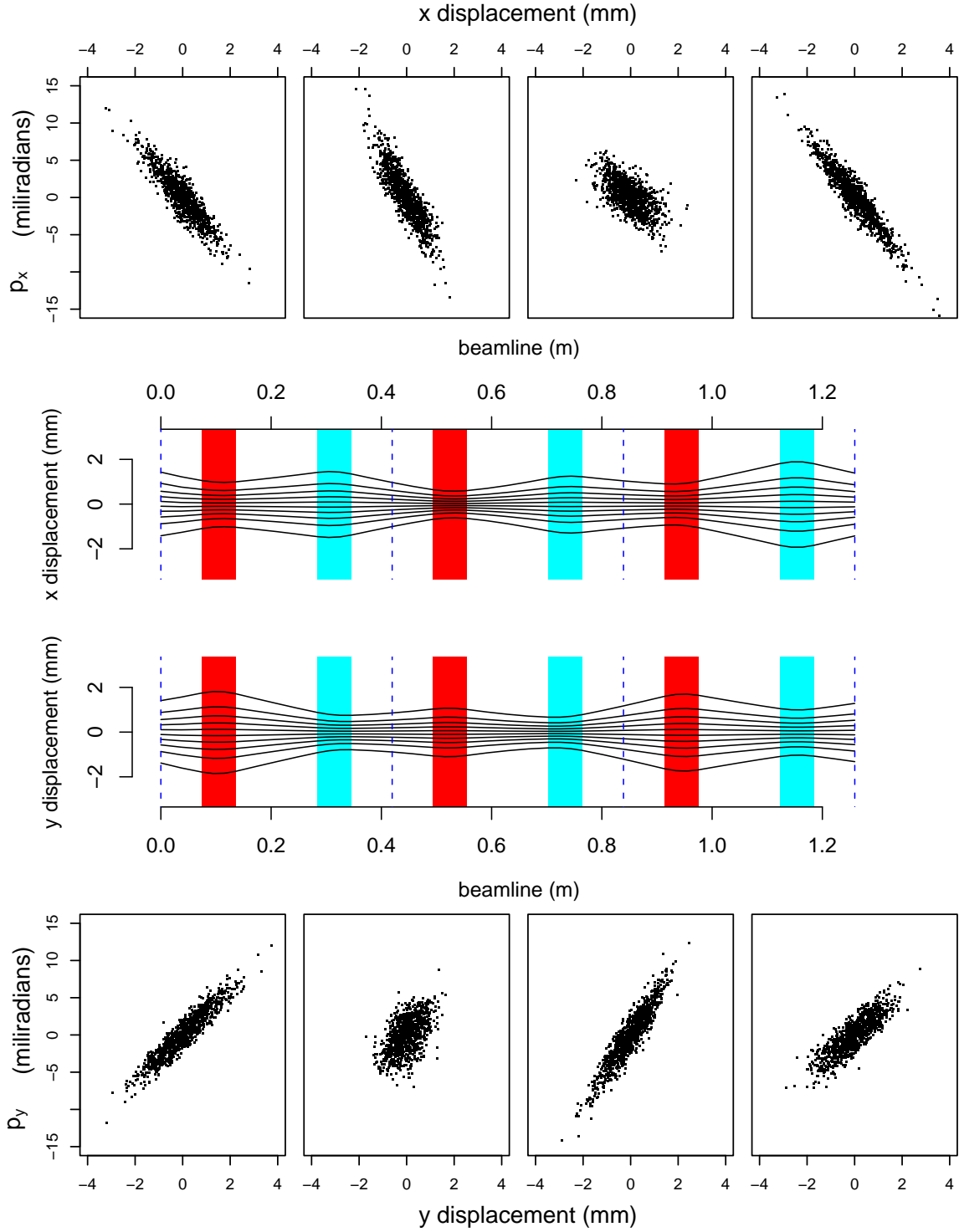
4

Figure 4: Particle simulation down the LEDA beamline. The central diagrams show part of the LEDA beamline consisting of three pairs of focusing and defocusing quadrupole magnets denoted by the dark and light shaded regions respectively. Wirescan locations are given by the four dotted vertical lines. At these four beamline locations, the resulting x- and y-phase space is shown for this particular simulation.

However, k_x^2 and k_y^2 do have a common magnitude that we take to be positive, i.e., $|k_x^2| = |k_y^2| = k^2$. For the LEDA case, $k_z = 0$. The equation of motion for the x direction is, in matrix form:

$$\mathbf{x}''(s) = \mathbf{A}(s)\mathbf{x}(s), \tag{5}$$

where

$$\mathbf{x}(s) = \begin{pmatrix} x(s) \\ x'(s) \end{pmatrix}, \quad \text{and} \quad \mathbf{A}(s) = \begin{pmatrix} 0 & 1 \\ -k_x^2 & 0 \end{pmatrix}, \tag{6}$$

with a similar equation for the y plane.

We can derive the lumped transfer matrix approximation used in this study as follows. Assuming that k_x, say, is approximately constant over a beam element of length l, we can integrate the x equation to get the solution:

$$\mathbf{x}(l) = \mathbf{\Phi}(l)\mathbf{x}(0), \tag{7}$$

where,

$$\mathbf{\Phi}(l) = \begin{pmatrix} \cos(kl) & \frac{1}{k}\sin(kl) \\ -k\sin(kl) & \cos(kl) \end{pmatrix} \equiv \mathbf{F}; \quad \mathbf{\Phi}(l) = \begin{pmatrix} \cosh(kl) & \frac{1}{k}\sinh(kl) \\ k\sinh(kl) & \cosh(kl) \end{pmatrix} \equiv \mathbf{D} \tag{8}$$

for the focusing and defocusing cases, respectively. $\mathbf{D}$ can be formally obtained from $\mathbf{F}$ by continuing k to purely imaginary values $k \rightarrow ik$. Similar solutions hold true for the y direction. For the z direction, the equation of motion integrates trivially to give the final type of transfer matrix:

$$\mathbf{O}(l) \equiv \begin{pmatrix} 1 & l \\ 0 & 1 \end{pmatrix} \tag{9}$$

Now to model the beam we simply treat each element as lumped and multiply the appropriate transfer matrices together in both the x and y directions. By construction, this procedure provides a linear map between the input phase space image and the phase space images at various points along the beam.

How valid is this approximate description of the beam? Of course, it is not possible to actually measure the phase space image at some beam point. Therefore, to address the issue of validating the approximate transfer model, we turn to a detailed numerical simulation of LEDA using the MaryLie/IMPACT accelerator code (Ryne, Qiang, Dragt, Habib, Mottershead, Neri, Samulyak and Walstrom (2003)). We set up the LEDA configuration in ML/I, providing two known input phase images to the code representing, respectively, simple and complex input distributions. The code
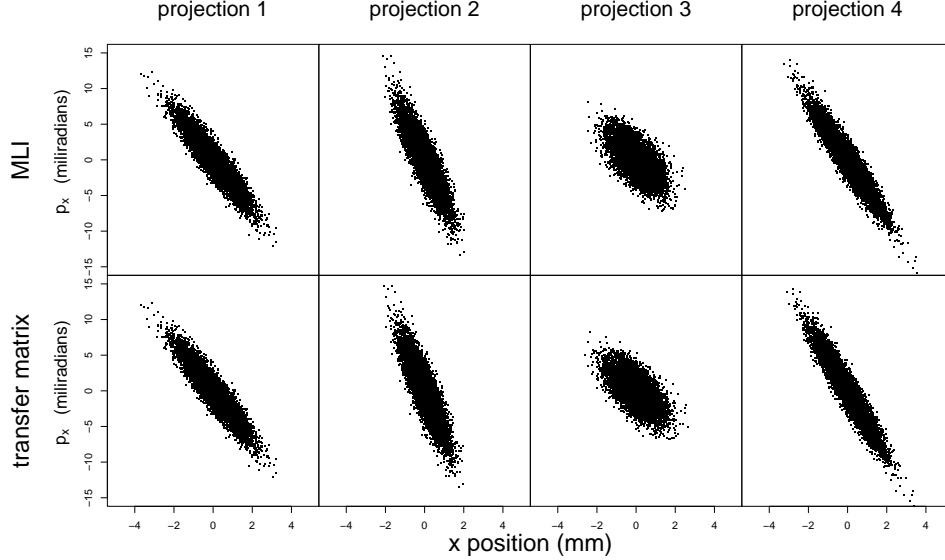
6

Figure 5: Comparison of transfer matrix (lower row) with MLI particle simulation (upper row) for a LEDA configuration.

then provides the phase space images at downstream points. As seen in Fig. 5 the downstream phase space images computed by the transfer matrix method compare favorably with those computed by ML/I for this application.

4 Statistical formulation

We start with an outline of our Bayesian reconstruction of the beam phase space images from the projection data. Each of the steps will be discussed in detail below. First, we parametrize the phase space image at the initial wirescan station using the method of process convolutions (Higdon (2002)). Given a candidate image at the initial scanner location, we can use the method of transfer matrices to propagate the phase space image down the beam to subsequent wirescan stations. At any given scanner location, we project the image in a linear fashion to generate a candidate beam profile to compare with the actual data. Therefore, given a parametrization of the first phase space image, and given the appropriate transfer matrices for the beam, we have a fitted value for the wirescan output that is used to construct the sampling model for the data.

To write down a posterior distribution on the parameters of the first phase space image, we need a prior distribution on the parameters governing the initial phase space image. The priors we use enforce a certain amount of smoothness on the image, in agreement with our physical intuition about the experiment. Finally, we explore the posterior on the initial image parameters using
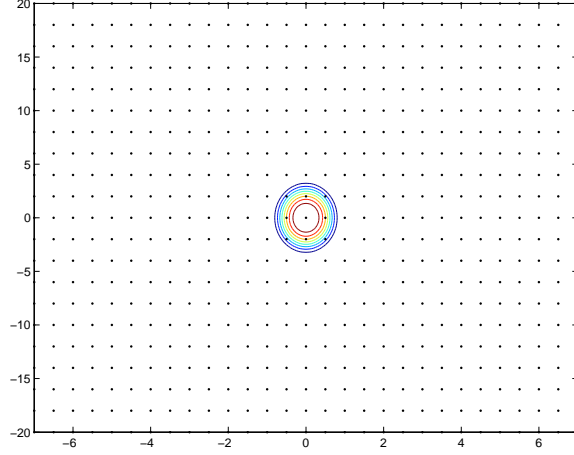
7

Figure 6: Grid of kernels used to parametrize initial phase space image.

Markov Chain Monte Carlo techniques (Besag, Green, Higdon and Mengersen (1995)), which allow us to not only estimate the mean posterior image but to assess the degree of variation the projection data allow.

4.1 Image representation

We represent the initial phase image as a process convolution (Higdon (2002)), using tricube kernels located on a grid of fixed centers, $\{x_i\}_{i=1}^{p}$, where p is the number of kernels chosen to represent the image. Each x_i is a point in the two dimensional image plane. The "standard" representation in the analysis below uses an $x \times p_x$ grid of $29 \times 21 = 609$ kernels to parametrize the image, as shown in Fig. 6.

The expansion of the image $z(x)$ is:

$$z(x) = \sum_{i=1}^{p} K^{(1)}(x - x_i)u_i, \tag{10}$$

where $u = (u_1, u_2, \ldots, u_p)^T$ is a vector of process parameters. The kernel $K^{(1)}$ has the tricube form:

$$K^{(1)}(x) \propto (1 - |x|^3/r^3)^3 I[|x| \leq r], \tag{11}$$

where r is the range of the kernel, and $I[\cdot]$ is the indicator function. This representation also provides good dimensionality reduction, an important computational benefit.

We remark that the number of initial image parameters u is almost always a good deal larger than the number of total wire scan data points. For example, a typical number of total wire scan

8

data points in the x direction, say, is 260. This is more than a factor of 2 less than a typical number of image parameters. Obviously, inference conducted in such a data sparse environment will strongly depend on the prior information brought to bear in the analysis.

4.2 Likelihood model

We now need to relate the initial phase image parameter vector to the projection data at all the scanners. Using products of transfer matrices to propagate the particle state vector (i.e., position and momentum) down the beamline is clearly a linear operation. In addition, projecting the image onto the horizontal and vertical axes is also a linear operation. As discussed above, the phase space image at the initial scanner is obtained from a vector of process parameters, u, by the action of a matrix $K^{(1)}$. We denote the projection matrix that takes the resulting image down to a beam profile by P_x or P_y, depending on the direction under consideration. Therefore, a candidate beam profile for the x direction at the initial scanner is:

$$\hat{y}^{(1)} = P_x K^{(1)} u. \tag{12}$$

If we denote the transfer matrix that sends the initial phase space image into the image at the second location by Φ_2, the beam profile at the second location is:

$$\hat{y}^{(2)} = P_x \Phi_2 K^{(1)} u, \tag{13}$$

and similarly for the succeeding locations. If we put $\Phi_1 = I$ we can gather all the scans into a single data structure:

$$\hat{y} = [\hat{y}^{(1)}; \hat{y}^{(2)}; \hat{y}^{(3)}; \hat{y}^{(4)}] = [P_x \Phi_1 K^{(1)} u; P_x \Phi_2 K^{(1)} u; P_x \Phi_3 K^{(1)} u; P_x \Phi_4 K^{(1)} u]. \tag{14}$$

For brevity of expression, we write all these scans together as $\hat{y}(u) = Ku$, where K is the matrix including the beam transfer matrices; the kernel representation; and the projection matrices.

To include experimental error, we model the data as:

$$y = Ku + \eta = \hat{y}(u),$$

with $\eta \sim N(0, \frac{1}{\lambda_e}I)$. We have denoted the experimental precision by $\lambda_e = 1/\sigma_e^2$.

The observed data are obtained from wirescans that convert proton "hits" into electrical current and are — to first order — scaled counts. Hence we use the standard square root transform to stabilize the variance, resulting in the likelihood function:

$$l(y|u, \lambda_e) \propto \lambda_e^{n/2} \exp(-\frac{\lambda_e}{2}(\sqrt{y} - \sqrt{Ku})^T(\sqrt{y} - \sqrt{Ku})), \tag{15}$$

9

where the precision parameter λ_e scales the variance, which is assumed to be constant on this transformed scale, and n is the total number of data points in all the scans. As seen below, we estimate λ_e in the MCMC. The assumption of independence in the error structure is our starting point; it is convenient, allows for straightforward estimation, and corresponds well to the experimentalists' beliefs about the experiment. As will be seen, this assumption leads to efficient reconstructions that account for most of the variation of the physical measurements. If, however, it emerges that correlations among errors prove to be significant, they can be accommodated in our analysis.

We note here that the mixture prior on u described in the following section ensures that $Ku \geq 0$. This restriction follows from the fundamental positivity of the projection data and the linearity of the data model. If we were to allow negative u's we could generate negative wirescans and possibly negative particle densities in the phase space image, which is physically impossible. However, vanishing u's are quite possible and indeed prominent in the phase image away from the image core.

4.3 Image prior

We construct the prior in steps. To address the positivity constraint on the u's, we use iid mixture priors for each u_i that put a finite mass on $u_i = 0$ and distribute the rest of the mass over $u_i > 0$ using a $N(0, 1/\lambda_u)$ density. Denote the mixture proportion by α. Then the prior density $\pi(u|\lambda_u)$ is given by:

$$\pi(u|\lambda_u) = \prod_{i=1}^{p} \left[\alpha\delta(u_i) + (1-\alpha)2\left(\frac{\lambda_u}{2\pi}\right)^{1/2} \exp\left(-\frac{\lambda_u}{2}u_i^2\right) I[u_i > 0] \right], \tag{16}$$

where $\delta(x)$ is the Dirac delta function, $I[\cdot]$ is the indicator function, as above, and λ_u controls the width of the half-normal mixture component.

We have found it sufficient to simply set λ_u to 1. Inspection of the prior for u reveals this choice allows sizable values of u to if needed. This reflects our physical knowledge of the phase image as being primarily composed of a prominent core surrounded by a halo. If λ_u is allowed to assume too large a value, the u's will be concentrated about 0 and will not attain the large values necessary to reconstruct the core of the image. We also set the mixture parameter α to 1/2. This choice amounts to adopting the prior assumption that a given point in the phase space image is as likely to be inside the core as outside the core.

Finally, we take a uniform prior over $\mathbb{R}^+$ for λ_e:

$$\pi(\lambda_e) \propto 1. \tag{17}$$

As we will describe below, this mixture prior leads to some slight complications in the MCMC algorithm.

4.4 Posterior estimation

Bayes' theorem gives the posterior as the product of the prior and the likelihood. The results given above give a joint posterior for u and λ_e:

$$\pi(u, \lambda_e | \lambda_u) \propto \pi(u|\lambda_u) \times \lambda_e^{n/2} \exp(-\frac{\lambda_e}{2}(\sqrt{y} - \sqrt{Ku})^T(\sqrt{y} - \sqrt{Ku})), \tag{18}$$

which allows us to recognize the full conditional density for λ_e as a gamma density:

$$\lambda_e | u, y \sim \Gamma(n/2 + 1, \frac{1}{2}(\sqrt{y} - \sqrt{Ku})^T(\sqrt{y} - \sqrt{Ku})). \tag{19}$$

We explore the joint posterior distribution using a single site update Metropolis-Hastings algorithm for the parameter vector u and a Gibbs step for the experimental precision λ_e (Besag et al. (1995)). In our code, we take the logarithm of the posterior distribution for numerical convenience. In the Metropolis-Hastings step, we tuned the proposal width to obtain an acceptance ratio of approximately $1/2$.

The update $u_i | u_{-i}, \lambda_e$ is as follows. First, we draw $u_i^{''}$ from a uniform proposal distribution with width r: $u_i^{''} \sim U[u_i - r, u_i + r]$. Then we set $u_i^{'} = 0$ if $u_i^{''} \leq 0$ and $u_i^{'} = u_i^{''}$ otherwise. We set the updated value $u_i^* = u_i^{'}$ with probability γ, where:

$$\gamma = \min\left(1, \frac{\pi\left(u_i^{'}|\lambda_u\right) q\left(u_i^{'} \rightarrow u_i\right)}{\pi\left(u_i|\lambda_u\right) q\left(u_i \rightarrow u_i^{'}\right)}\right), \tag{20}$$

and the proposal density q is given by:

$$q(x \rightarrow y) = \begin{pmatrix} \frac{r-x}{2r} & \text{if } x > 0 \text{ and } y = 0 \\ \frac{1}{2r} & \text{if } x = 0 \text{ and } y > 0 \\ \frac{1}{2r} & \text{if } x > 0 \text{ and } y > 0 \\ \frac{1}{2} & \text{if } x = 0 \text{ and } y = 0 \end{pmatrix}. \tag{21}$$

5 Results

The computed mean reconstructions in both the x and y directions for scans $20 - 26$ are shown in Fig. 7. Under the mean images are plots of the wirescan data (red dots), the 95% pointwise
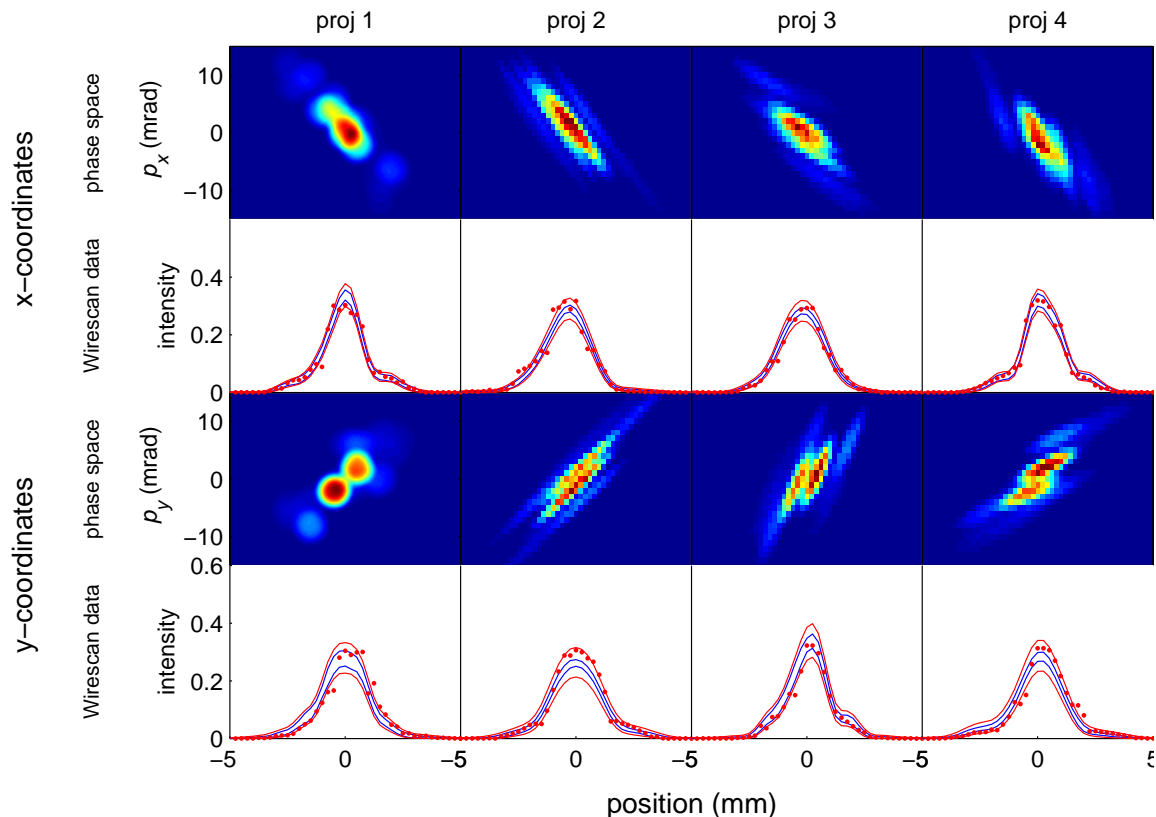
Figure 7: Reconstructions for scans 20–26. Red dots: wirescan data. Blue lines: 95% pointwise confidence limits for mean scan. Red lines: 90% pointwise confidence limits for a predicted wirescan.

confidence limits for the mean wirescan (blue lines) and the 90% pointwise confidence limits for a predicted experimental wirescan (red lines). Fig. 8 is a histogram of λ_e for the x direction showing the posterior estimate of the experimental precision. Similarly, the computed mean reconstructions for scans $40 - 45$ are shown in Fig. 9.

We performed several checks on the robustness of our scheme. Numerical experimentation indicated that about 1000 iterations, discarding 50 for burn-in, were sufficient to achieve robust reconstructions. For comparison, Fig. 10 shows the reconstruction for a 5000 iteration run.

To check our sensitivity to the resolution of the kernel grid, we also computed a finely zoned reconstruction (1539 kernels). The result for scans 20–26 is shown in Fig. 11.

Comparing Figs. 11 and 7 we see that the coarse-scale reconstruction effectively imposes a higher degree of smoothness on the image. The fine-scale reconstruction tends to follow the data too closely and, of course, is much more computationally intensive.
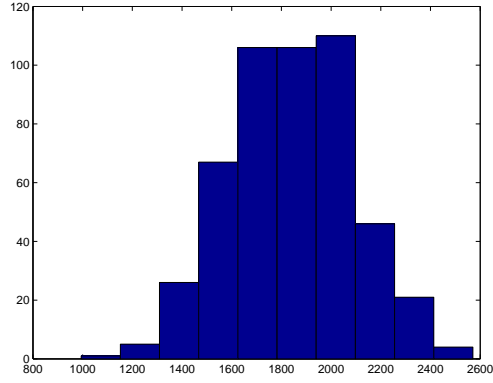
12

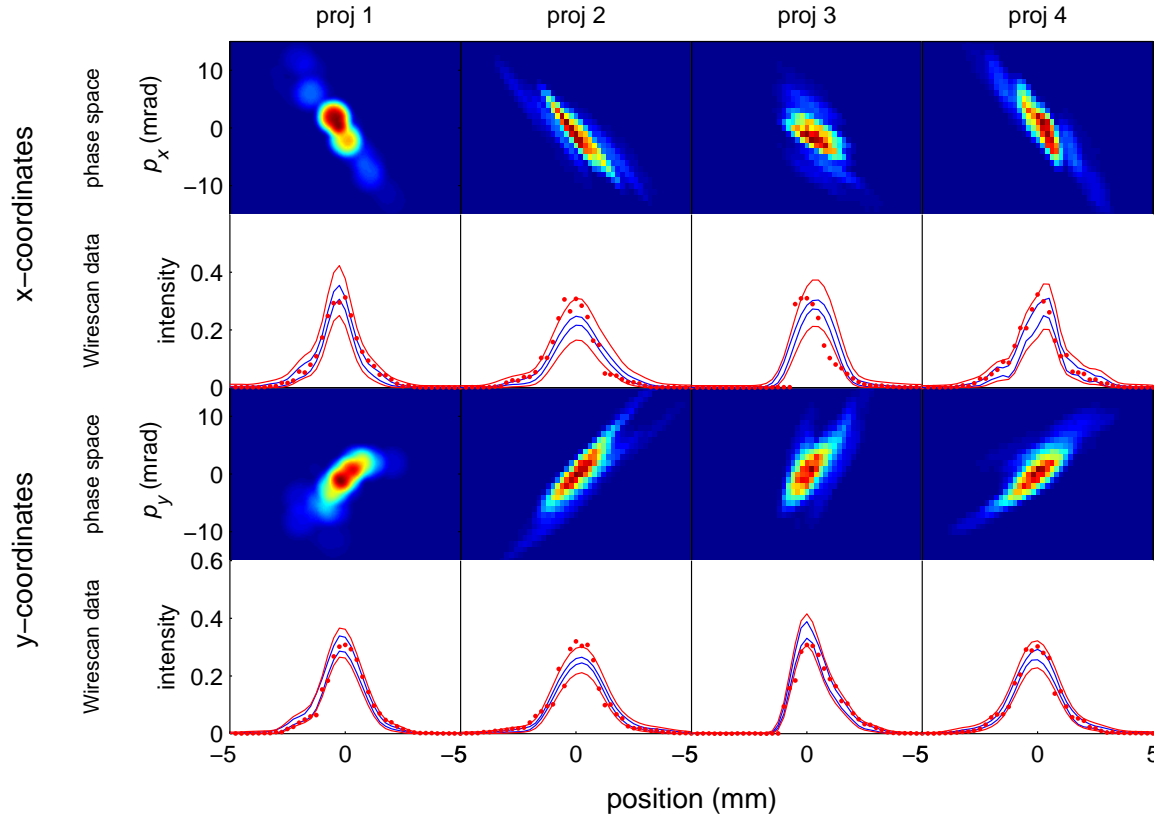Figure 8: Posterior distribution of experimental precision.

Figure 9: Reconstructions for scans 45–51. Red dots: wirescan data. Blue lines: 95% pointwise confidence limits for mean scan. Red lines: 90% pointwise confidence limits for a predicted wirescan.
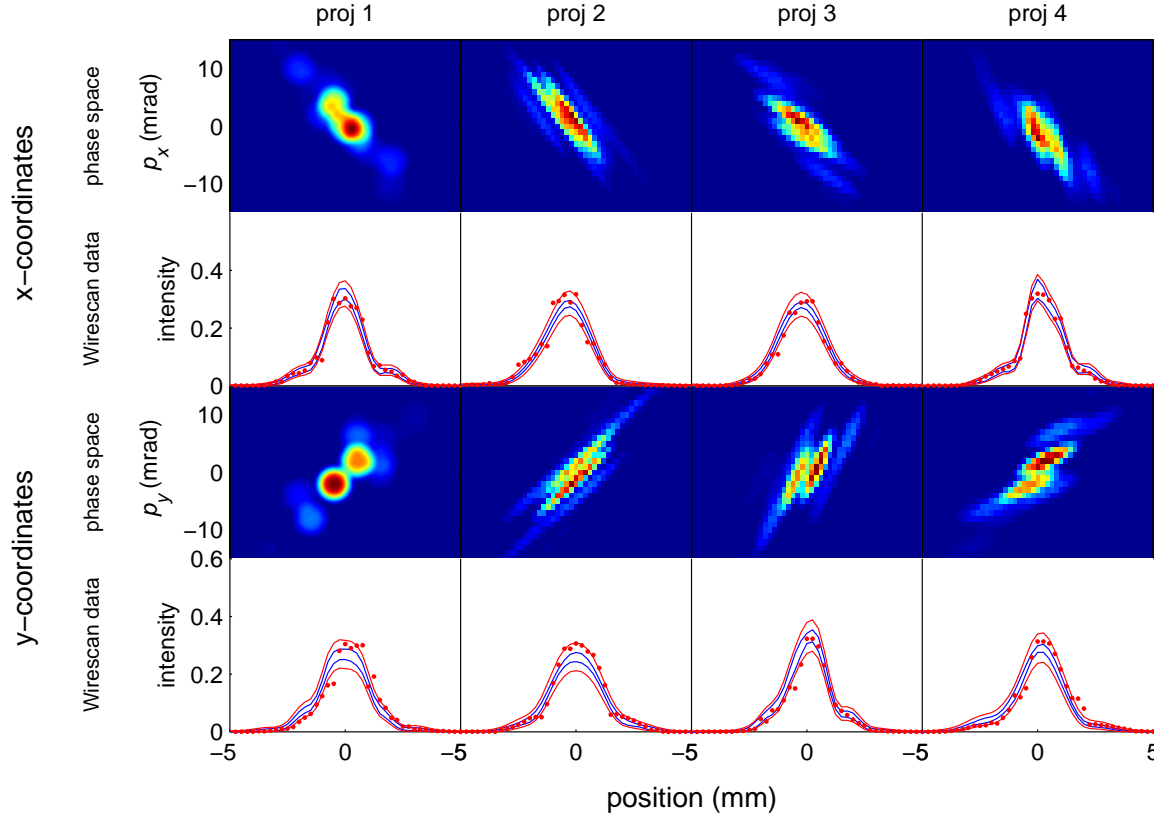
13

Figure 10: 5000 iteration reconstruction for scans 20–26. Red dots: wirescan data. Blue lines: 95% pointwise confidence limits for mean scan. Red lines: 90% pointwise confidence limits for a predicted wirescan.
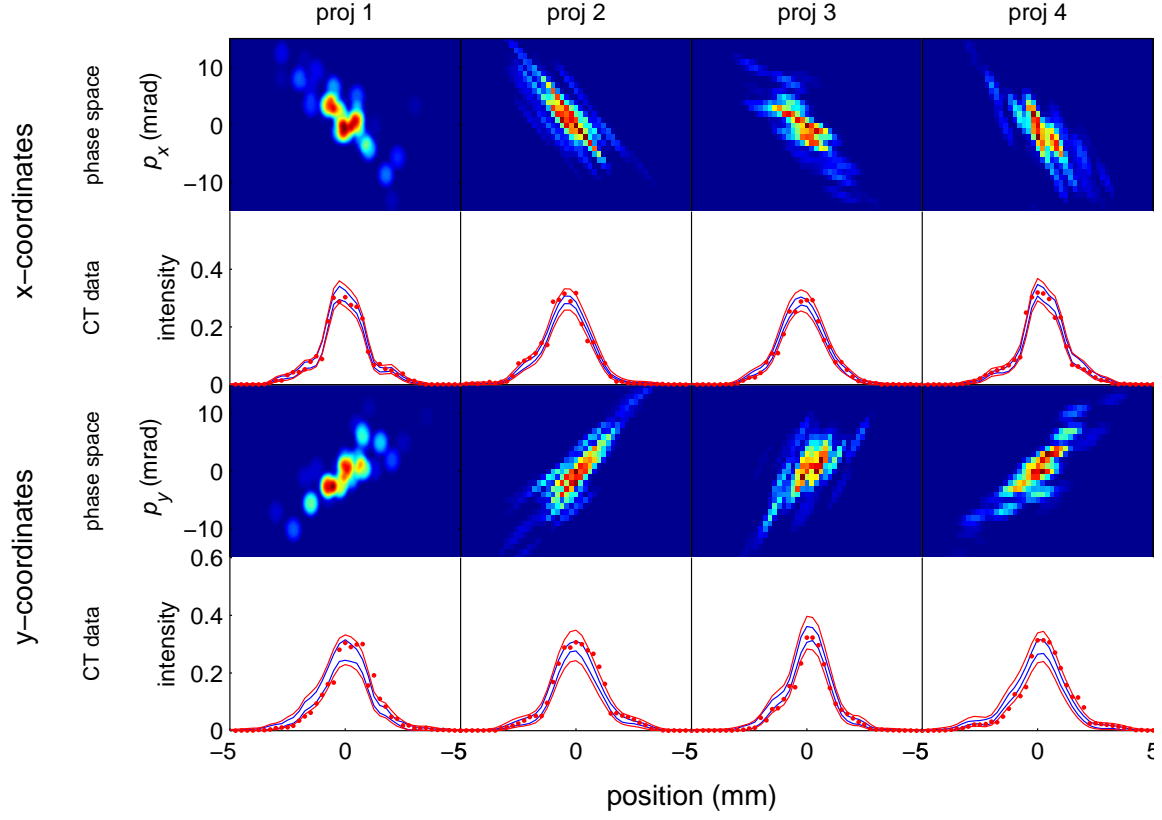
Figure 11: Fine-scale reconstructions for scans 20–26. Red dots: wirescan data. Blue lines: 95% pointwise confidence limits for mean scan. Red lines: 90% pointwise confidence limits for a predicted wirescan.

6 Discussion

Reconstruction of particle beam phase space images from projection data provides an excellent application of modern computational Bayesian inference. The quality of the reconstructions is quite good, and convincingly demonstrates the ability of our method to extract physically sensible phase information from the projection data. The ability to capture complex phase information outside the image cores is particularly noteworthy. This type of phase image detail has been previously unavailable to accelerator physicists, and should enhance their ability to design efficient accelerators. On the statistical side, the physical constraint that the image parameters be nonnegative leads to a complex prior structure, accompanied by nontrivial complications in the MCMC update.

The major limiting approximation made in this paper is the use of the transfer matrix method to model the beam dynamics. As we have shown, this approximation is satisfactory for the LEDA accelerator. However, this approximation may not suffice for future accelerator designs. If not, then the transfer matrices used here will need to be replaced by full numerical simulations like ML/I. In this case, computational expense will become an important factor. MCMC techniques will place a high premium on finding methods of efficiently modeling the simulator output. The development of such methods for for accelerator codes like ML/I is a topic for future research.

References

Allen, C. K. and Pattengale, N. D. (2002). Theory and technique of beam envelope simulation: Simulation of bunched particle beams with ellipsoidal symmetry and linear space charge forces, *Technical Report LA-UR-02-4979*, Los Alamos National Laboratory.

Allen, C. K., Chan, K. C. D., Colestock, P. L., Crandall, K. R., Garnett, R. W., Gilpatrick, J. D., Lysenko, W., Qiang, J., Schneider, J. D., Schulze, M. E., Sheffield, R. L., Smith, H. V. and Wangler, T. P. (2002). Beam-halo measurements in high-current proton beams, *Physical Review Letters* **89**: 214802–1–214802–4.

Besag, J., Green, P. J., Higdon, D. M. and Mengersen, K. (1995). Bayesian computation and stochastic systems (with discussion), *Statistical Science* **10**: 3–66.

Higdon, D. (2002). Space and space-time modeling using process convolutions, *in* C. Anderson, V. Barnett, P. C. Chatwin and A. H. El-Shaarawi (eds), *Quantitative Methods for Current Environmental Issues*, Springer Verlag, London, pp. 37–56.

Ryne, R., Qiang, J., Dragt, A., Habib, S., Mottershead, T., Neri, F., Samulyak, R. and Walstrom, P. (2003). MaryLie/IMPACT: A parallel beam dynamics code with space charge, *Proceedings of the Computational Accelerator Physics Conference, Oct. 15-18, 2002*.